



## Major TCR Repertoire Perturbation by Immunodominant HLA-B\*44:03-Restricted CMV-Specific T Cells

Attaf, Meriem; Malik, Amna; Severinsen, Mai C; Roider, Julia; Ogongo, Paul; Buus, Søren; Ndung'u, Thumbi; Leslie, Alasdair; Kløverpris, Henrik N; Matthews, Philippa C; Sewell, Andrew K; Goulder, Philip

*Published in:*  
Frontiers in Immunology

*DOI:*  
[10.3389/fimmu.2018.02539](https://doi.org/10.3389/fimmu.2018.02539)

*Publication date:*  
2018

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY](#)

*Citation for published version (APA):*  
Attaf, M., Malik, A., Severinsen, M. C., Roider, J., Ogongo, P., Buus, S., Ndung'u, T., Leslie, A., Kløverpris, H. N., Matthews, P. C., Sewell, A. K., & Goulder, P. (2018). Major TCR Repertoire Perturbation by Immunodominant HLA-B\*44:03-Restricted CMV-Specific T Cells. *Frontiers in Immunology*, 9, [2539]. <https://doi.org/10.3389/fimmu.2018.02539>



# Major TCR Repertoire Perturbation by Immunodominant HLA-B\*44:03-Restricted CMV-Specific T Cells

Meriem Attaf<sup>1,2†</sup>, Amna Malik<sup>3†</sup>, Mai C. Severinsen<sup>4,5</sup>, Julia Roider<sup>4,6</sup>, Paul Ogongo<sup>4,7,8</sup>, Søren Buus<sup>5</sup>, Thumbi Ndung'u<sup>4,9</sup>, Alasdair Leslie<sup>4,10</sup>, Henrik N. Kløverpris<sup>4,5,9</sup>, Philippa C. Matthews<sup>11</sup>, Andrew K. Sewell<sup>1,2</sup> and Philip Goulder<sup>3\*</sup>

## OPEN ACCESS

### Edited by:

Remy Bosselut,  
National Cancer Institute (NCI),  
United States

### Reviewed by:

Paul Austin Moss,  
University of Birmingham,  
United Kingdom  
Eric Huseby,  
University of Massachusetts Boston,  
United States

### \*Correspondence:

Philip Goulder  
philip.goulder@paediatrics.ox.ac.uk

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
T Cell Biology,  
a section of the journal  
Frontiers in Immunology

**Received:** 25 June 2018

**Accepted:** 15 October 2018

**Published:** 14 November 2018

### Citation:

Attaf M, Malik A, Severinsen MC,  
Roider J, Ogongo P, Buus S,  
Ndung'u T, Leslie A, Kløverpris HN,  
Matthews PC, Sewell AK and  
Goulder P (2018) Major TCR  
Repertoire Perturbation by  
Immunodominant  
HLA-B\*44:03-Restricted  
CMV-Specific T Cells.  
Front. Immunol. 9:2539.  
doi: 10.3389/fimmu.2018.02539

<sup>1</sup> Division of Infection and Immunity, Cardiff University School of Medicine, Cardiff, United Kingdom, <sup>2</sup> Systems Immunity Research Institute, Cardiff University, Cardiff, United Kingdom, <sup>3</sup> Department of Paediatrics, University of Oxford, Oxford, United Kingdom, <sup>4</sup> Africa Health Research Institute, Durban, South Africa, <sup>5</sup> Laboratory of Experimental Immunology, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark, <sup>6</sup> Department of infectious diseases, Medizinische Klinik IV, Ludwig-Maximilians-University Munich, Munich, Germany, <sup>7</sup> School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, South Africa, <sup>8</sup> Department of Tropical and Infectious Diseases, Institute of Primate Research, Nairobi, Kenya, <sup>9</sup> Department of Infection and Immunity, University College London, London, United Kingdom, <sup>10</sup> Department of Infection and Immunity, University College London, London, United Kingdom, <sup>11</sup> Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

Lack of disease during chronic human cytomegalovirus (CMV) infection depends on the maintenance of a high-frequency CMV-specific T cell response. The composition of the T cell receptor (TCR) repertoire underlying this response remains poorly characterised, especially within African populations in which CMV is endemic from infancy. Here we focus on the immunodominant CD8+ T cell response to the immediate-early 2 (IE-2)-derived epitope NEGVKAAW (NW8) restricted by HLA-B\*44:03, a highly prevalent response in African populations, which in some subjects represents >10% of the circulating CD8+ T cells. Using pMHC multimer staining and sorting of NW8-specific T cells, the TCR repertoire raised against NW8 was characterised here using high-throughput sequencing in 20 HLA-B\*44:03 subjects. We found that the CD8+ T cell repertoire raised in response to NW8 was highly skewed and featured preferential use of a restricted set of V and J gene segments. Furthermore, as often seen in immunity against ancient viruses like CMV and Epstein-Barr virus (EBV), the response was strongly dominated by identical TCR sequences shared by multiple individuals, or “public” TCRs. Finally, we describe a pair “superdominant” TCR clonotypes, which were germline or nearly germline-encoded and produced at remarkably high frequencies in certain individuals, with a single CMV-specific clonotype representing up to 17% of all CD8+ T cells. Given the magnitude of the NW8 response, we propose that this major skewing of CMV-specific immunity leads to massive perturbations in the overall TCR repertoire in HLA-B\*44:03 individuals.

**Keywords:** T cell, T cell receptor, cytomegalovirus, HLA-B\*44:03, repertoire, sequencing

## INTRODUCTION

Human cytomegalovirus (CMV) is a highly prevalent  $\beta$ -herpesvirus (1, 2). Although CMV infection can lead to severe clinical complications in immunocompromised patients such as in the setting of AIDS (3–5), in transplant patients on immunosuppressive regimens (6, 7) or following congenital infection (8), it is usually asymptomatic. CD8<sup>+</sup> T cells play a central part of protection against CMV disease (9, 10). Adoptive transfer of CMV-specific CD8<sup>+</sup> T cells into immunodeficient recipients effectively restores antigen-specific CD8<sup>+</sup> T cell immunity, reduces CMV viraemia and prevents disease (9, 10). Conversely, lack of CD8<sup>+</sup> T cell-mediated responses correlates with CMV replication and disease progression. The frequency of CMV-specific CD8<sup>+</sup> T cells following natural infection is unusually high, constituting up to 10% of the CD8<sup>+</sup> T cell memory compartment in peripheral blood in adults (11). These responses are maintained for decades and, in the case of certain specificities, may increase in frequency even further over time through a process termed “memory inflation” (12, 13).

Whether the size of the CMV-specific T cell response impacts the ability of the immune system to mount effective responses against other pathogens is unclear, but it is apparent that HIV disease progression is substantially more rapid in the case of HIV-CMV coinfection (14, 15). Age is also likely to be factor, and the contribution of CMV to immunosenescence and increased mortality from infectious disease in the elderly is still a topic of active debate (16, 17). Because the human T cell receptor (TCR) repertoire is finite and limited to ~25 million TCRs (18), it is possible that in individuals in whom 10% or more of CD8<sup>+</sup> T cell responses are dedicated to control of CMV infection, the availability of optimal TCRs to control infections other than CMV might be constrained. This hypothesis remains untested, as high-throughput sequencing studies of TCR repertoires following CMV infection are relatively few, especially in populations where CMV is endemic from infancy (19, 20). In sub-Saharan Africa, most children become CMV-infected, in the first 3 months of life and CMV infection rates approach 90% in infants reaching 12 months of age (1, 21). Thus, understanding the molecular basis of CMV-specific T cellular responses has the potential to affect the lives of millions, in population groups that are particularly vulnerable.

We recently characterised an immunodominant CMV-specific response in African populations that is restricted by HLA-B\*44:03 (22). This class I molecule is expressed in ~15% of African populations (23) and differs by a single amino acid from the closely related HLA-B\*44:02 (24), the most prevalent HLA-B\*44 subtype in Caucasian populations. Among individuals expressing HLA-B\*44:03, an unusually high fraction of CD8<sup>+</sup> T cells (up to 17%) is specific for the immediate-early-2 (IE-2<sub>463–473</sub>)-derived epitope NEGVKAAW (hereafter referred to as NW8). The disproportionate magnitude of the NW8-specific response in B\*44:03+ individuals is likely to be underpinned by strong biases in the TCR repertoire, leading to preferential selection of NW8-specific CD8<sup>+</sup> T cells against this HLA background, although this possibility has yet to be investigated.

Using high-throughput sequencing, we sought to characterise the TCR repertoire of peptide-MHC tetramer-sorted NW8-specific CD8<sup>+</sup> T cells from 20 HLA-B\*44:03+ CMV-infected individuals. We found that the TCR repertoire of NW8-specific CD8<sup>+</sup> T cells displayed limited clonal diversity and usage of a restricted set of V and J genes. The response was characterised by the presence of identical or “public” TCRs in multiple unrelated individuals. Public TCR sequences were often dominant within an individual and characterised by central amino acid motifs which were found to be either entirely germline-encoded (in the case of TCR- $\alpha$  chains), or near germline ( $\beta$  chains). We conclude that the high-magnitude NW8 response causes a major perturbation of the TCR repertoire, with disproportionate expansion of germline clonotypes that are highly shared amongst HLA-B\*44:03+ individuals.

## MATERIALS AND METHODS

### Study Subjects

We studied 20 HIV-infected ART-naïve adult subjects from a previously described cohort in Durban (South Africa) (25). The subjects studied had a mean age of 23 years (IQR, 21–31 years), a mean absolute CD4<sup>+</sup> T cell count of 477/mm<sup>3</sup> (IQR, 205–107,500) and a median HIV viral load of 20,500 copies/ml (IQR 816–23,000). HIV viral load was measured using the Roche Amplicor version 1.5 assay according the manufacturer's instructions; CD4<sup>+</sup> T cell counts were measured by flow cytometry. Samples utilised for global TCR repertoire sequencing from genomic DNA were obtained from subjects of similar ethnic background and also recruited in Durban, South Africa. Ethics approval, for both studies, was given by the KwaZulu-Natal Review Board (Durban cohort). All subjects provided written informed consent.

### Cell Sorting

Peptide-MHC tetramers were generated as previously described (26). Cryopreserved PBMC (1 million per staining condition) from the recipients were stained with PE-conjugated or APC-conjugated peptide-MHC tetramers, anti-CD3 Pacific Orange (Invitrogen, UK), anti-CD8 V450 (BD Biosciences, Oxford, UK) antibodies and near-IR Live/Dead marker (Invitrogen, Paisley, UK). The samples were sorted on BD FACSAria (BD Biosciences, Paisley, UK) and tetramer+ CD8<sup>+</sup> cells were collected in 350  $\mu$ L of RLT lysis buffer (QIAGEN, Hilden, Germany).

### TCR Sequencing

RNA extraction was carried out using the RNEasy Micro kit (Qiagen, Hilden, Germany) as previously described (27). Briefly, cDNA was synthesized using the 5'/3' SMARTer kit (Takara Bio, Paris, France) according to the manufacturer's instructions. The SMARTer approach used a Murine Moloney Leukaemia Virus (MMLV) reverse transcriptase, a 3' oligo-dT primer and a 5' oligonucleotide to generate cDNA templates flanked by a known, universal anchor sequence. A reverse primer specific for the TCR- $\alpha$  or the TCR- $\beta$  constant region (C $\alpha$ R1 5' CCATAGACCTCATGTCTAGCACAG-3' or C $\beta$ R1 5'-GAGACCCCTCAGGCGGCTGCTC-3', Eurofins Genomics,

Ebersberg, Germany) was then used together with an anchor-specific forward primer (Takara Bio, Paris, France) in the following PCR reaction: 2.5  $\mu$ L template cDNA, 0.25  $\mu$ L High Fidelity Phusion Taq polymerase, 10  $\mu$ L 5x Phusion buffer, 0.5  $\mu$ L DMSO (all from Thermo Fisher Scientific, UK), 1  $\mu$ L dNTP (50 mM each, Life Technologies, Paisley, UK), 1  $\mu$ L of each primer (10  $\mu$ M), and nuclease-free water for a final reaction volume of 50  $\mu$ L. Subsequently, 2.5  $\mu$ L of the first PCR products were used to set up a nested PCR as above, using nested primers (C $\alpha$ R2 5'-GGTGAATAGGCAGACAGACTTGTC-3' or C $\beta$ R2 5'-TGTGTGGCCAGGCACACCAGTGTG-3', Eurofins Genomics, Ebersberg, Germany and anchor-specific primer from Takara Bio, Paris, France). For both PCR reactions, cycling conditions were as follows: 5 min at 94°C, 30 cycles of 30 s at 94°C, 30 s at 63°C, 90 s at 72°C, and a final 10 min at 72°C. The final PCR products were loaded on a 1% agarose gel and purified with the QIAEX II gel extraction kit (Qiagen, Hilden, Germany). Purified products were barcoded, pooled and sequenced on an Illumina MiSeq instrument using the MiSeq v2 reagent kit (Illumina, Cambridge, UK).

TCR sequencing from genomic samples from whole blood was carried out by Adaptive Biotechnologies, using the Immunoseq platform, as previously described (28).

## TCR Sequence Analysis

Analysis of the raw TCR sequences was performed using MiXCR (29). MiXCR employs a built-in library of reference germline V, D, J, and C gene loci from the ImMunoGeneTics (IMGT) database (imgt.org). The IMGT nomenclature for TCR gene segments was used throughout the study. All aligned, in-frame, antigen-specific TCR clonotypes are available in the VDJDB repository at vjdb.cdr3.net (30). By convention, the TCR third complementarity-determining region (CDR3) is written from the cardinal Cys residue to the conserved, J-encoded Phe residue.

## Statistical Analysis

### Repertoire Clonality (Evenness)

TCR repertoire clonality is given by the Shannon evenness index ( $J'$ ):

$$J' = \frac{H'}{\ln(n)} \quad H' = -\sum_{i=1}^n p_i \ln(p_i)$$

Where  $p_i$  is the frequency of the  $i^{\text{th}}$  clonotype in a population of  $n$  clonotypes.

$J'$  is undefined for monoclonal samples. Low  $J'$  values approaching 0 indicate minimal evenness as seen after clonal expansion of antigen-specific species. The maximal  $J'$  value is 1, when all clonotypes have equal frequencies, i.e., the population is perfectly even.

### Repertoire Compositional Similarity

Compositional similarity between TCR repertoires were assessed using the Morisita-Horn similarity index ( $C_{MH}$ ):

$$C_{MH} = \frac{2 \sum_{i=1}^c f_i g_i}{\sum_{i=1}^c (f_i^2 + g_i^2)}$$

where  $f_i = n_{1i}/N_1$  and  $g_i = n_{2i}/N_2$ ,  $n_{1i}$  and  $n_{2i}$  are the number of copies of the  $i^{\text{th}}$  clonotype in samples 1 and 2, and  $N_1$  and  $N_2$  are the total number of TCRs in samples 1 and 2, respectively. The summations in the numerator and the denominator are over the unique clonotypes ( $c$ ) in both samples. The similarity indices range in value from 0 (minimal similarity; the samples are entirely different) to 1 (maximal similarity; the samples are identical). The  $C_{MH}$  similarity index accounts for both the number of common clonotypes and the distribution of clone sizes and is sensitive to the clone sizes of the dominant clonotypes.  $C_{MH}$  similarity calculations were performed using the “numpy” package in Python.

## Statistical and Graphical Analysis

All pairwise statistical tests were performed in Prism v7.0 (GraphPad, San Diego, USA) unless stated otherwise. Strength of association between two variables was analysed by Spearman's rank test.  $P$  values < 0.05 were considered significant.

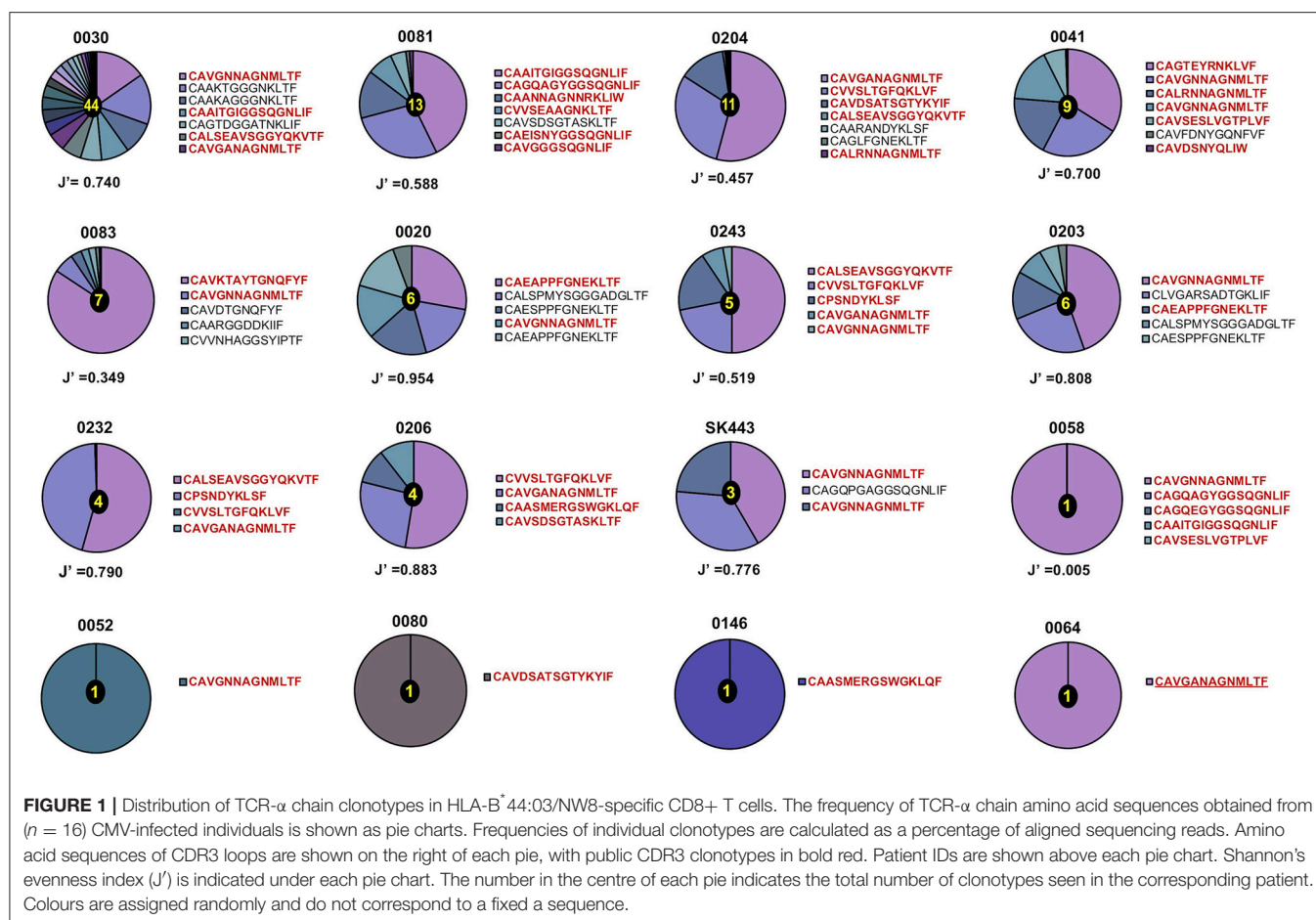
## RESULTS

### NW8-Specific T Cells Display Limited TCR Diversity

We sequenced the TCR- $\alpha$  and TCR- $\beta$  chain repertoires of NW8-specific CD8+ T cells sorted from 20 HLA-B\*44:03+ individuals from Durban, South Africa. The mean size of the tetramer+ population was 3.88% of CD3+CD8+ T cells (IQR 0.36-6.05) and the mean number of cells sorted for sequencing was 10,169 cells (IQR 327-11,791) (Supplementary Table 1 and Supplementary Figure 1). In total, 1,750,000 reads from TCR- $\alpha$  chain samples and 700,000 from TCR- $\beta$  chains were generated using the Illumina MiSeq platform from these samples. This translated into 335,891 functional, in-frame TCR- $\alpha$  chain sequences and 91,154 TCR- $\beta$  chains.

A total of 53 TCR- $\alpha$  chain clonotypes were identified in 16 individuals and 51 TCR- $\beta$  clonotypes in 18 individuals (Figures 1, 2). TCR repertoire richness, as measured by the number of clonotypes per patient varied widely across the cohort. This was particularly apparent in the case of TCR- $\alpha$  chains, for which the number of clonotypes ranged from 1 to 42. However, TCR- $\beta$  chain samples were more homogeneous in size and the total numbers of TCR- $\beta$  clonotypes only varied between 1 and 11. Although intuitively this could be attributed to variation in the number of sorted cells, the number of clonotypes did not correlate with the size of the tetramer+ population or with the number of cells sequenced (Supplementary Table 1). TCR clonality was evaluated using the Shannon evenness index ( $J'$ ).  $J'$  is undefined (approaching zero) when the sample is monoclonal, low when the distribution of clonotypes is uneven, and 1 when all clonotypes have the same frequency (See Materials and Methods). Four out of 16 TCR- $\alpha$  and five out of 18 TCR- $\beta$  repertoires were strictly monoclonal (Figures 1, 2). Of particular interest was patient 0064 whose TCR- $\alpha$  and TCR- $\beta$  repertoires were both monoclonal, which by default indicated that the NW8-specific response in this individual comprised of only one TCR. Other repertoires showed evidence of preferential expansion of a limited set of clonotypes, as suggested by low  $J'$  values. To





confirm this extreme oligoclonality was the result of bona fide antigen-specific expansion and not methodological bias, TCR-α and TCR-β chain samples were amplified using an independent primer set, cloned into a plasmid vector and sequenced. Using this alternative approach, we found that the clonotypic hierarchy was preserved, confirming the limited TCR diversity in NW8-specific response as genuine (**Supplementary Figure 2**). Thus, both the TCR-α and the TCR-β repertoires of NW8-specific CD8+ T cells displayed evidence of clonal expansion, which in some cases led to dominance of a single clone, or very few clones at best.

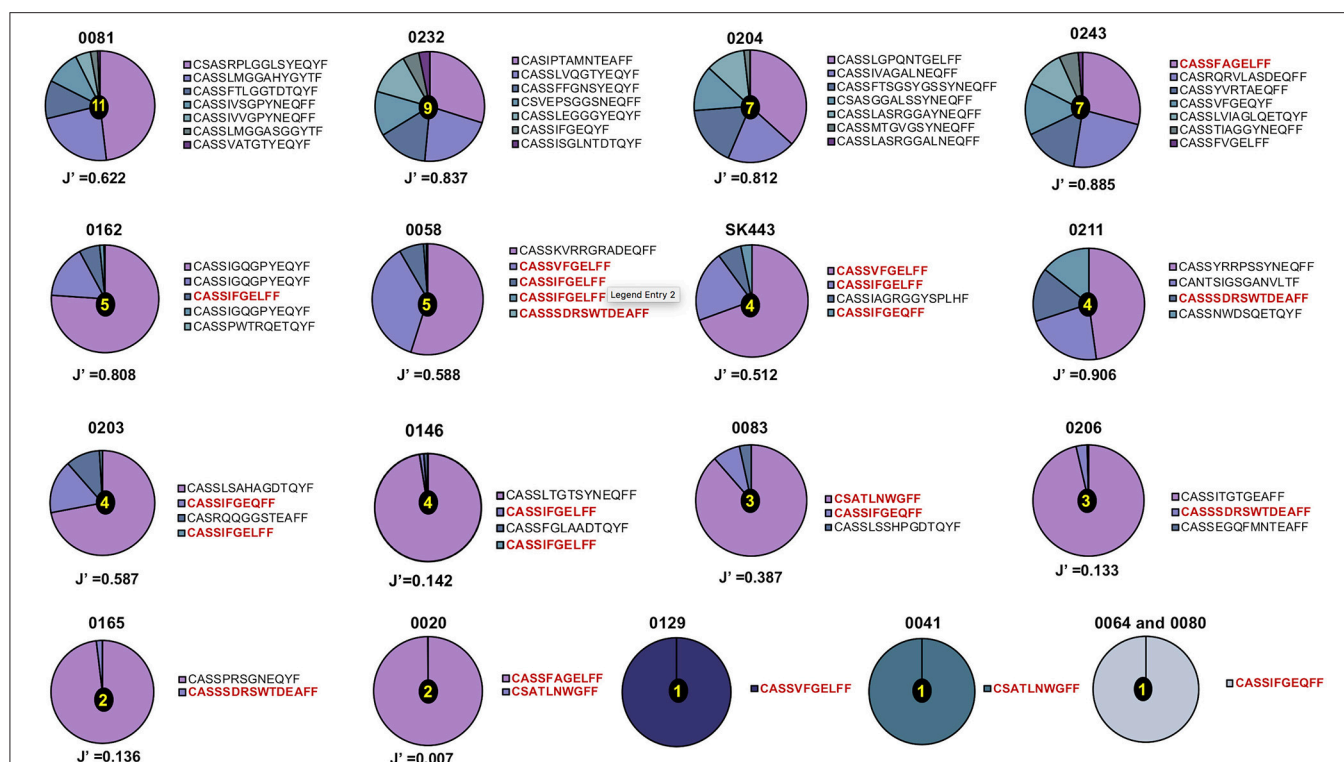
## The NW8-Specific CD8+ T Cell Repertoire Is Characterised by Restricted V and J Segment Usage

TCRs are generated early in T cell ontogeny by recombination of V (D, in the case of TCR-β) and J gene segments (31). Antigen-driven selection of TCR clonotypes often leads to skewed distributions of V and J genes (32–35). Because this phenomenon is particularly apparent in the setting of human viral infection, we sought to determine whether NW8-driven clonal selection would also lead to a narrowing of V and J gene usage. We looked at the distribution of TRAV and TRBV genes, first across individuals, then as a proportion of all sequences obtained in the study. NW8-specific CD8+ T cells indeed displayed strongly skewed

gene usage (**Figure 3**). Out of the 16 individuals sequenced, 14 (87%) harboured at least one TCR encoded by the TRAV20 gene segment. Moreover, almost a third (32%) of all TCR-α chains were encoded by the TRAV20 gene segment (**Figure 3A**). Similarly, the TRBV19 gene segment was found in 15 out of the 18 individuals sequenced (83%), and 43% of all TCR-β sequences were encoded by TRBV19 (**Figure 3B**). Accordingly, certain TRAV-TRAJ gene combinations were more frequent than others. Among TCR-α chain repertoires, the TRAV20-TRAJ39 pair was the most frequent (**Figure 3C**). TCR-β sequences showed an even more striking bias for TRBV19 pairing with TRBJ2-2, 2-1 and 2-7 (**Figure 3D**). Thus, the NW8-specific CD8+ T cell repertoire is characterised by restricted V and J segment usage. Altogether, the preferential selection of the TRAV20 and TRBV19 genes across individuals with a small set of TRAJ/TRBJ (respectively) is evidence for antigen-driven selection of public clonotypes (further discussed in the next section).

## High-Frequency TCR Clonotypes Are Highly Shared Amongst HLA-B\*44:03+, CMV-Infected Individuals

Viral infection such as CMV often leads to the selection of predictable repertoires with extensive sharing of TCR clonotypes (31, 36). As mentioned above, the selection of a narrow set of V and J genes pointed to the public nature of the NW8. To



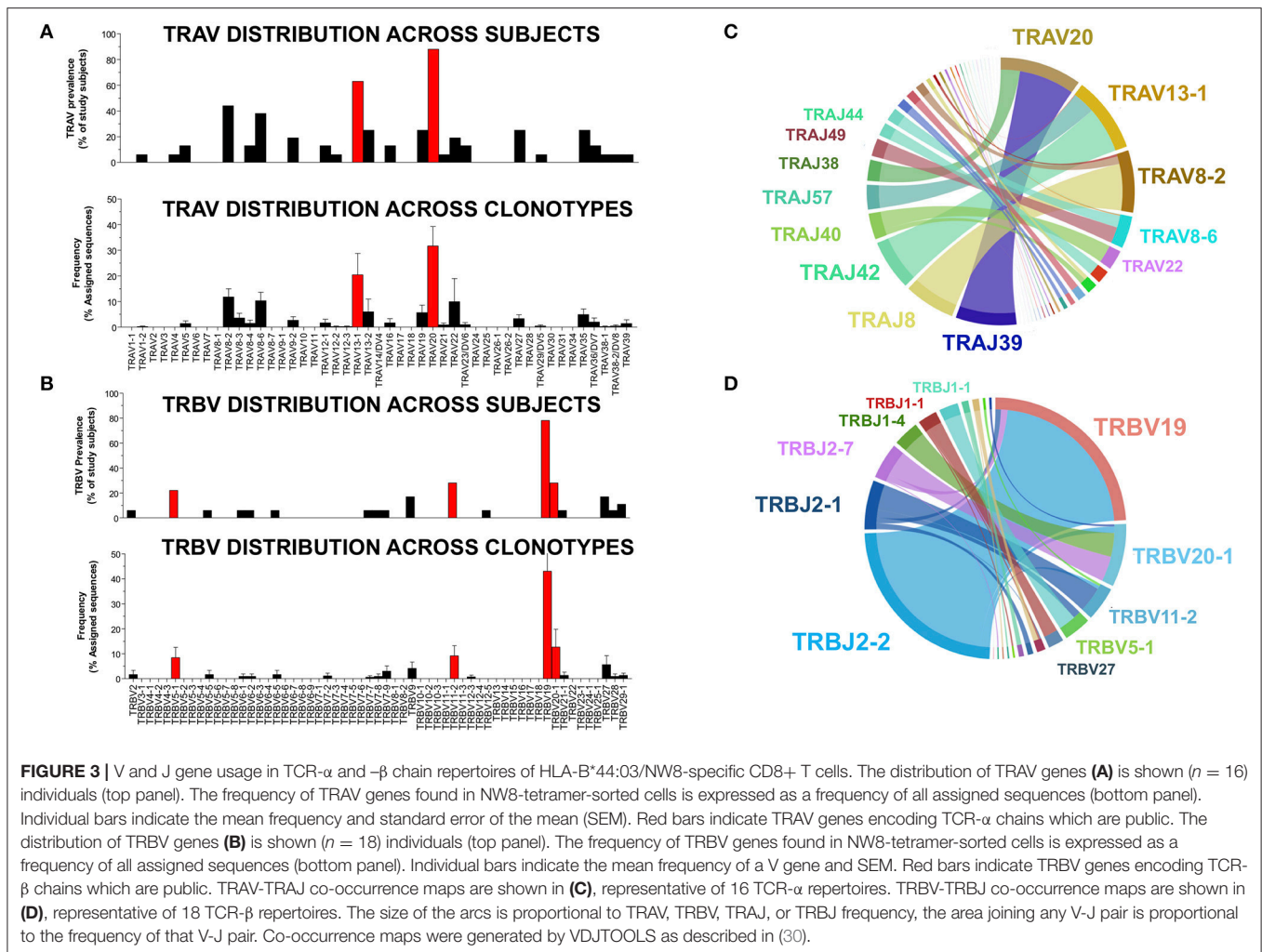
**FIGURE 2 |** Distribution of TCR-β chain clonotypes in HLA-B\*44:03/NW8-specific CD8+ T cells. The frequency distribution of TCR-β chain amino acid sequences obtained from ( $n = 18$ ) CMV-infected individuals is shown as pie charts. Frequencies of individual clonotypes are calculated as a percentage of aligned sequencing reads. Amino acid sequences of CDR3 loops are shown on the right of each pie, with public CDR3 clonotypes in bold red. Patient IDs are shown above each pie chart. Shannon's evenness index ( $J'$ ) is shown under each pie chart. The number in the centre of each pie indicates the total number of clonotypes seen in the corresponding patient. Colours are assigned randomly and do not correspond to a fixed a sequence.

test whether NW8-specific CD8+ T cells were indeed public, we used the Morisita-Horn similarity-index ( $C_{MH}$ ) to perform pairwise comparisons between every combination of two donors, as a measure of compositional similarity, or overlap (Figure 4).  $C_{MH}$  values lie between 0 (non-overlapping populations) and 1 (identical populations). The  $C_{MH}$  value for every combination of two donors, representing a total of 120 pairwise combinations for the TCR-α chain and 153 for TCR-β (see Materials and Methods). The mean  $C_{MH}$  value of the TCR-α chain repertoire was  $0.20 \pm 0.02$  (Figure 4A), indicating that on average, the patients in this study shared a fifth of their NW8-specific, TCR-α repertoire. Despite strong selection for TRBV19+ clonotypes, the mean  $C_{MH}$  of TCR-β repertoires was only  $0.09 \pm 0.10$  (Figure 4B). The lower  $C_{MH}$  in TCR-β chain repertoires is likely to merely reflect the inclusion of a D segment in TCR-β transcripts, leading to more junctional diversity, compared to TCR-α. Nonetheless, multiple public TCR chains were identified, both in TCR-α and TCR-β repertoires (Figures 1, 2, 4).

Out of 53 TCR-α chain clonotypes, 32 were public. Two related clonotypes, CAVGNNAGNMLTF and CAVGANAGNMLTF, both TRAV20/TRAJ39 as anticipated and differing by just a single amino acid, were the two most abundant of all clones sequenced (30.4 and 3.7%,

respectively). These clones were also the most extensively shared. CAVGNNAGNMLTF was found in 12/16 (75%) of subjects studied and CAVGANAGNMLTF in 7/16 (44%). To confirm this extreme bias, TCR-α chains were sequenced using the Immunoseq platform from whole blood genomic DNA of 26 individuals of the (28) same ethnicity, four of which were HLA-B\*44:03+ (Supplementary Figure 3). The frequency of the CAVGANAGMLTF and CAVGNNAGMLTF clonotypes cumulated to 3.3, 1.1, and 2.3% of total productive TCRs (derived from both CD4+ and CD8+ T cells), in three of the four HLA-B\*44:03+ subjects. By contrast, in HLA-B\*44:03-negative subjects, the frequency of these clonotypes did not exceed 0.01%. This demonstrates, using a different methodology, that the superdominant clonotypes CAVGANAGMLTF and CAVGNNAGMLTF are highly enriched in HLA-B\*44:03+ individuals, and expanded to comparably high frequencies as observed by tetramer staining.

Similarly, the most abundant public TCR-β clonotypes, CASSIFGEQYF and CASSIFGELFF (TRBV19/TRBJ2-1 and TRBV19/TRBJ2-2, respectively), were shared by 7/18 (39%) of individuals studied. Consistent with previous observations (37), the higher the frequency of a given clonotype, the greater the likelihood that this would be a public TCR sequence. This correlation between clonotype frequency and the degree



of publicity was statistically significant for TCR- $\alpha$  chains (Spearman's  $r = 0.4290$ ,  $p = 0.0015$ ; **Figure 4C**) but not for TCR- $\beta$  chains (**Figure 4D**). This suggests that the previously described relationship between clonal abundance and publicity is weak in the case of the NW8 response and that other factors contribute to “superdominance” of certain TCR clonotypes in HLA-B\*44:03+ individuals.

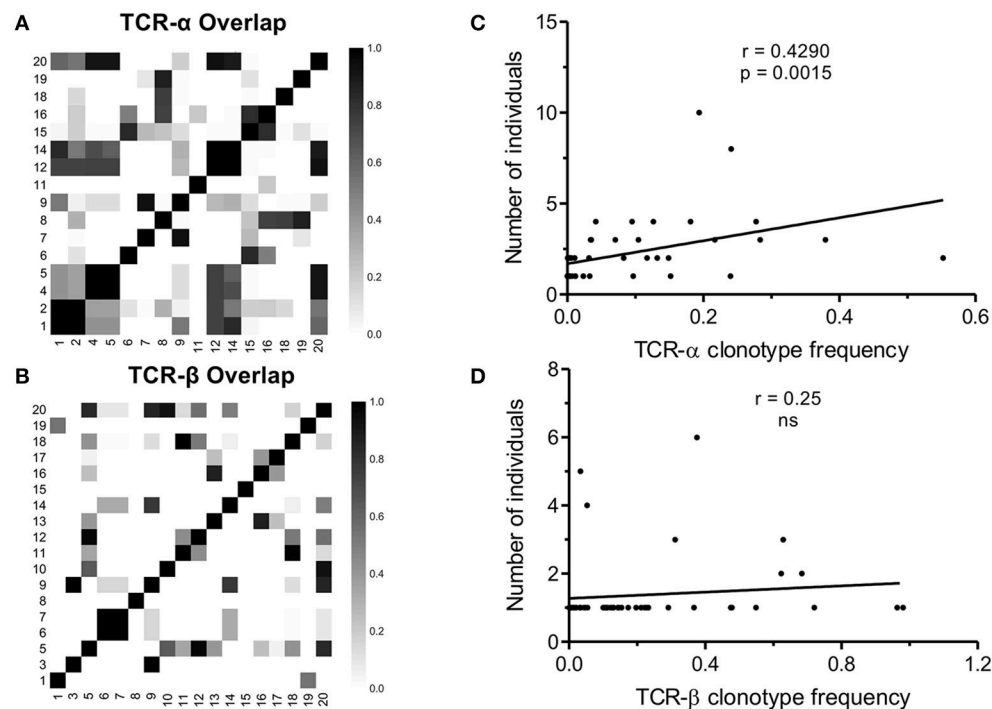
The public amino acid sequences were often encoded by more than one distinct nucleotide sequence (**Supplementary Tables 2, 3** and **Supplementary Figure 4**). For example, the most prevalent public TCR- $\beta$  clonotype, CASSIFGEQFE, was encoded by four different nucleotide sequences. Similarly, the most frequently occurring TCR- $\alpha$  public clonotype, CAVGNNAGNMLTF, was also encoded by four different nucleotide sequences. Clonotypes that occurred in a greater number of individuals tended to also be encoded by a greater number of distinct nucleotide sequences. This is often referred to as convergent recombination, whereby redundancy in the genetic code leads to generation of TCR clonotypes with degenerate nucleotide sequences at high frequency (38, 39). For TCR- $\beta$  chains, this association was statistically significant (Spearman's  $r = 0.798$ ,  $p < 0.0001$ , **Supplementary Figure 4**).

Overall the high prevalence of public TCRs in NW8-specific CD8 $^{+}$  T cells, that can be encoded by several nucleotide sequences, is further evidence of the strong selection for particular TCR amino acid sequences by the HLA-B\*44:03-NW8 peptide-MHC complex. The public nature of the NW8 response is likely to be underpinned by the conservation of structural features allowing the selection of “superdominant,” highly prevalent TCR chains against this MHC background.

### Public CDR3 Motifs Are Strictly Conserved Amongst HLA-B\*44:03+, CMV-Infected Individuals

TCR repertoire bias often includes the conservation of amino acid motifs within CDR3 loops (33, 34, 40, 41). Here we observed that several TCR- $\alpha$  chain clonotypes contained the sequence motif Asn-Ala-Gly (referred to as NAG thereafter) or Gly-Gly-Ser (GGS). TCR- $\alpha$  chains with the NAG motif were shared across more individuals than those with no motif ( $p < 0.0001$ ) or with the GGS motif ( $p = 0.0271$ ; **Figure 5A**). This also held true for TCR- $\beta$  chains containing the Ile-Phe-Gly (IFG) motif (**Figure 5B**). All TCR chains with a conserved motif were public.





**FIGURE 4 |** TCR sharing amongst NW8-specific CD8+ T cells from CMV-infected individuals. Compositional similarity amongst (A) TCR-α and (B) TCR-β chains was assessed using the Morisita-Horn ( $C_{MH}$ ) index and illustrated as a heat map. Patient numbers are shown along the x and y-axes.  $C_{MH}$  values range from 0 (no overlap) to 1 (perfect overlap). The average frequency of a given clonotype is plotted against the number of individuals sharing that clonotype, for (C) TCR-α chains ( $n = 53$ ) and (D) TCR-β chains ( $n = 51$ ). Spearman's rho and  $p$ -values are shown above each graph.

Conversely, private TCR chains displayed no apparent motif. The NAG motif was entirely encoded by the TRAJ region, specifically the TRAJ39 segment in 93.5% of clonotypes. Similarly, the GGS motif was largely encoded by TRAJ42. Strikingly, V gene usage was not conserved. NAG clonotypes were encoded by three different TRAV genes and GGS by as many as seven TRAV genes (Figure 5C). The use of several TRAV genes likely excludes the possibility of a germline-encoded interaction between TCR-α CDR1 and -2 loops with MHC, as has been observed with other epitopes (42, 43), and suggests a more prominent role for CDR3 loops in NW8 recognition.

TCR-β chains clonotypes with the IFG motif were strictly encoded by TRBV19, while TRBJ usage was less conserved (Figure 5D). Again, all IFG motif-containing chains were public. The central phenylalanine residue at position 6 was not germline-encoded in all cases and resulted from two different codons. The conservation of a non-germline residue again argues for structurally-imposed, antigen-driven selection of clonotypes bearing this amino acid at this position. Overall, we conclude that TCR publicity in the NW8-specific response is governed by structural constraints selecting for germline or near-germline encoded CDR3 loops.

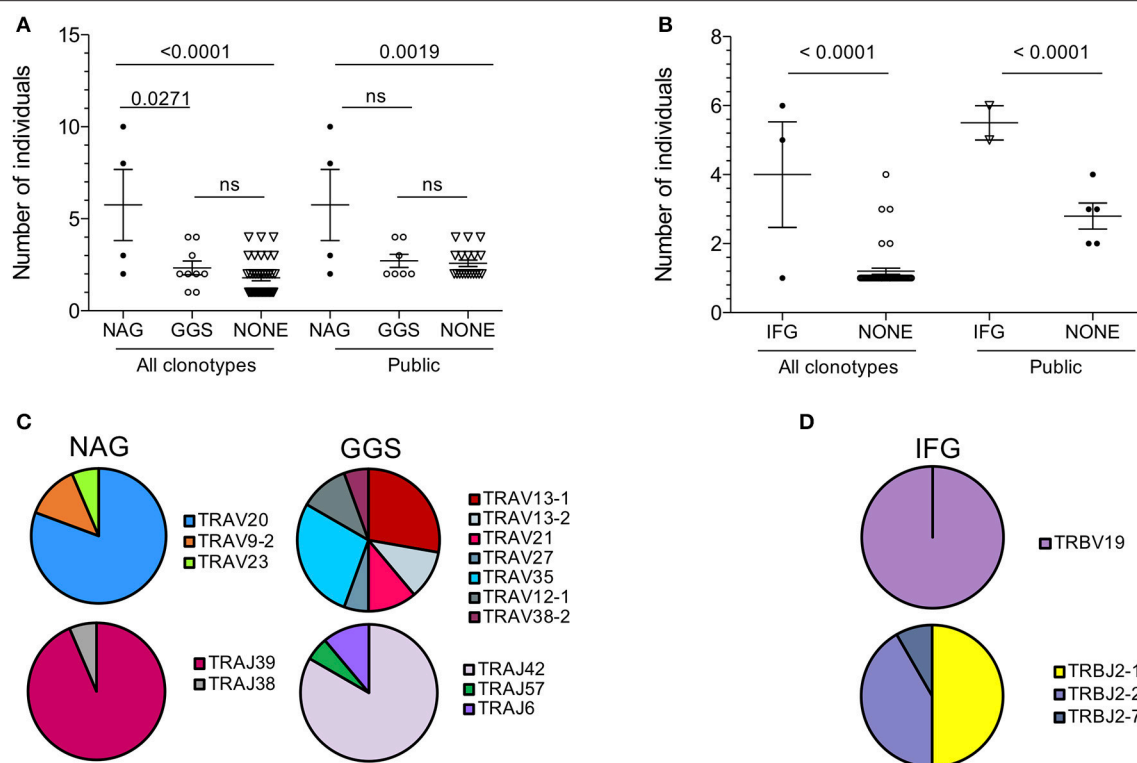
## DISCUSSION

In this study we sought to characterise the TCR repertoire of CD8+ T cells responding to an immunodominant CMV epitope in HLA B\*44:03+ individuals. Using high-throughput

sequencing, we dissected the NW8-specific TCR-α and -β chain repertoires of 20 CMV-infected subjects. These repertoires were found to be highly skewed, as indicated by preferential use of V and J genes, extensive TCR sharing and conservation of CDR3 motifs. TCR diversity and the specificity it underlies are the distinguishing feature of αβ T cells. Diversity in T cell responses has long been thought of as a strict requirement for the provision of protective immunity and full immune coverage. Nonetheless, it is now well established that human viral infections can give rise to narrow and skewed repertoires, often associated with TCR sharing in different individuals. Thus, the repertoire is often shaped toward a single or focused solution for the recognition of peptide-HLA complexes. Reports of TCR publicity in immune responses have accumulated in the literature, such that TCR sharing across individuals can now almost be taken as a rule, rather than the exception (36, 41, 44, 45).

The occurrence of public TCRs is nearly  $10^4$  times more likely than expected if all rearrangements were equiprobable, irrespective of HLA type (18). This is because public TCRs have distinct features, such as limited length and P/N diversification, which allows them to be generated frequently by the recombination machinery. A TCR sequence will have a higher likelihood of being produced if it is encoded by a nucleotide sequence with few additions and deletions. Redundancy in the genetic code also explains why TCR chains containing residues encoded by multiple codons (such as Gly or Ala) are very commonly produced and can occur at high frequency (38, 39). TCR production frequency has already been described





**FIGURE 5 |** Amino acid motif conservation in public clonotypes from HLA-B\*44:03/NW8-specific CD8+ T cells. **(A)** The number of individuals sharing NAG or GGS-containing TCR-α chains is shown. **(B)** The number of individuals sharing IFG-containing TCR-β chains is shown. **(C)** TRAV and TRAJ gene usage in NAG or GGS-containing TCR-α chains is illustrated as pie charts. Frequencies are calculated as fraction of NAG or GGS-containing amino acid sequences. **(D)** TRBV and TRBJ gene usage in IFG-containing TCR-β chains is illustrated as pie charts. Frequencies are calculated as fraction of IFG-containing amino acid sequences. Horizontal bars represent the mean and standard of the mean. Statistical significance is shown where  $p < 0.05$  (ns = not significant). Pie charts illustrate the frequency of a V or J gene. Colours are assigned randomly and do not correspond to a fixed a sequence.

as an important determinant of TCR sharing (37). Here, the public chains described in this study were nearly entirely germline-encoded or involved very limited changes at the CDR3 junction. This was effectively mirrored by the conservation of amino acid motifs. The TCR-α chain “NAG” motif described here was found to be entirely encoded by the J segment, and the “IFG” motif found in TCR-β chains only contained one added amino acid. The germline nature of these public chains again highlights the importance of TCR generation efficiency. In the case of NAG TCR-α chains, the conservation of the J segment, but not the V segment, excludes a potential role for germline CDR1 and CDR2 in recognition of HLA-B\*44:03 (46). Rather, this emphasises the starring role of the CDR3 loop in antigen recognition and structural constraints favouring the selection of public clonotypes with these conserved motifs (32–34).

The other distinguishing feature of NW8-specific repertoires was their extremely limited diversity. In some individuals, TCR-α and TCR-β chain repertoires were even monoclonal. This was notably the case of patient 0064, with a tetramer response of 16.8% (equivalent to 31,480 sorted cells) composed of a single TCR, CAVGANAGMLTF-CASSIFGEQFF. This, to our knowledge, is the largest reported T cell response mounted by

a single public T cell clone in humans, which initially led us to envisage a positive relationship between TCR publicity and the magnitude of the T cell response. However, we found that this was not the case, as the number of public TCRs did not correlate with the percentage of tetramer-positive cells or with the number of sorted cells. Instead, the frequency of a given clonotype positively correlated with the extent to which it is shared (i.e., the number of individuals harbouring that clonotype). Altogether, these results suggest that publicity in T cell responses does not necessarily equate to superior immunity. Rather, the dominance of public TCRs within an individual and across individuals can be explained by their high precursor frequency, or in other words, by a numerical competitive advantage (47).

The patient cohort recruited for this study were HIV-1 co-infected and receiving ART and it is well-established that such individuals have heightened CMV-specific T cell responses. Whether the expansion of superdominant clonotypes stems from dysregulated and heightened responses resulting from HIV-1 co-infection remains to be tested. It has been suggested that the increased frequency of CMV-specific T cells in ART patients may be due to abnormal responses to otherwise normal levels of subclinical CMV replication (48).

Similarly, the disproportionate representation of superdominant clonotypes could be another manifestation of HIV-mediated immune dysregulation.

The unusual magnitude of the HLA-B\*44:03-restricted NW8-specific CD8+ T cell response deserves comment. In most individuals, >3% of CD8+ T cells, and in one individual almost 1 in 5 CD8+ T cells (19%), are specific for this single response (22). Even by the standards of CMV-specific CD8+ T cell responses, which are renowned to be very large, this is an exceptional example. The progressive, prolonged expansion of these responses has been termed memory “inflation” (12, 49–51). It has been suggested that these accumulations of CMV-specific CD8+ T cells with age may be responsible for shortening of the human life-span as a result of excessive loss of naïve T cell numbers (52). If this is the case, it might be predicted that individuals with HLA-B\*44:03 might be at particular risk, not only of reduced ability to counter infections in later life, but also to combat coinfections such as HIV which also appear to induce large T cell expansions (53).

How this so-called memory “inflation” is established upon infection remains unclear, although it is possible that certain pathogens, like CMV, would adapt to host immunity to abrogate the contraction phase of the immune response. The composition of the “inflated” memory pool both at the cellular and clonotypic levels has also remained elusive (54). Our results are consistent with the notion that, in HLA-B\*44:03 individuals, the CD8+ T cell response to the NW8 epitope is associated with remarkably large clonotypic expansions reminiscent of memory inflation. In the absence of longitudinal TCR analysis, this hypothesis remains to be tested formally. Nonetheless, it is clear that this repertoire is a very limited repertoire of “superdominant clones” that in the extreme can occupy almost 20% of the CD8+ T cell pool. Whether this feature is typical of memory inflation in general, or specific to CMV, will have to be determined.

Intuitively, that highly prevalent pathogens such as CMV are best dealt with by TCRs which are highly prevalent across a population appears to be a beneficial trait. The *mhc* locus is the most gene-dense and the most polymorphic genetic complex known in jawed vertebrates, reflecting strong evolutionary pressure from highly mutable pathogens (55). Polymorphism at the *mhc* affects an individual's immune status by determining the collection of peptides made available to T cells during development and shaping of the TCR repertoire and also upon antigenic challenge (56). “Cat and mouse” evolution between the host immune system and invading pathogens is a well-documented phenomenon, particularly in the case of rapidly mutating pathogens such as retroviruses. By contrast, the high prevalence and genetic stability of herpesviruses such as CMV may lead to the conservation of a TCR genotype that is widely shared across the population as a result of frequent generation and selection of clonotypes which are germline or near-germline-encoded during T cell development (41, 57). Whether the “superdominant,” public TCRs identified in this study represent a built-in bank of receptors which can readily be mobilised against ubiquitous pathogens like herpesviruses remains to be determined. Understanding the molecular rules which govern how TCRs are generated and selected to seed the

periphery will be crucial in our understanding of public T cell responses.

## AUTHOR CONTRIBUTIONS

The study was conceived by AS and PG and funded by grants to AS and PG. MA, AM, MS, JR, and PO did the experimental work that was supervised by TN, SB, HK, PM, AS, and PG. The data were analyzed by MA, AM, HK, AS, and PG. The manuscript was written by MA, AM, AL, HK, PM, AS, and PG.

## FUNDING

This work was supported by the Wellcome Trust (grant to PG WT104748MA). AS is a Wellcome Senior Investigator (WT086716MA).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2018.02539/full#supplementary-material>

**Supplementary Figure 1 |** Tetramer-based sorting of CMV-specific CD8+ T cells. **(A)** tetramer+ populations are shown in flow cytometry plots, with HLA-B\*44:03-SL9 (derived from HIV-p24) tetramer along the y-axis and HLA-B\*44:03-NW8 tetramer along the x-axis. The plots fail to show double-positive cell populations, indicating specificity of staining. Patient identification number are shown in red within each plot. **(B)** FACS plots showing the gate boundaries set for sorting of CD8+HLA-B\*44:03-NW8 tetramer+ cells (labelled as P2, orange). The gates were set on Live CD3+CD8+ cells. Patient identification number are shown in red within each plot.

**Supplementary Figure 2 |** Sanger sequencing confirms clonal dominance in TCR- $\alpha$  and TCR- $\beta$  chain repertoires. The frequency of TCR- $\alpha$  (top) and TCR- $\beta$  chain (bottom) clonotypes is shown for four patients. Pie charts on the left-hand side represent data obtained using high-throughput sequencing, on the right-hand side Sanger sequencing.

**Supplementary Figure 3 |** Superdominant clonotypes are readily identifiable in HLA-B\*44:03-positive individuals. TCR- $\alpha$  chain sequencing was carried out on genomic DNA derived from whole blood, using the “ImmunoSeq platform” available at Adaptive Biotechnologies (28). The cumulative frequency of the superdominant clonotypes CAVGANAGMLTF and CAVGNAGMLTF is shown for  $n = 26$  individuals, of which four are HLA-B\*44:03+ (red bars).

**Supplementary Figure 4 |** Number of nucleotide sequences and TCR publicty. The number of nucleotide sequences encoding a given clonotype is plotted against the number of individuals sharing that clonotype, for **(A)** TCR- $\alpha$  chains and **(B)** TCR- $\beta$  chains. Spearman's rho and  $p$ -values are shown above each graph.

**Supplementary Table 1 |** Patient HLA-B type and size of tetramer-specific response. Four-digit HLA-B types are given for each patient, along with the size of the tetramer-specific response. Tetramer+ cells were gated on live CD3+CD8+ T cells.

**Supplementary Table 2 |** Public TCR clonotypes are encoded by multiple nucleotide sequences. The nucleotide and amino acid sequence of superdominant public clonotypes is shown, together with their absolute read count and relative frequency.

**Supplementary Table 3 |** TCR- $\alpha$  chain clonotypes are shared amongst HLA-B\*44:03+ individuals and encoded by multiple nucleotide sequences. TCR- $\alpha$  chain sequencing from genomic DNA was carried out in three HLA-B\*44:03+ individuals, using the ImmunoSeq platform (Adaptive Biotechnologies). Nucleotide and amino acid sequences are shown, together with their absolute count and relative frequency.

## REFERENCES

- Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Rev Med Virol.* (2010) 20:202–13. doi: 10.1002/rmv.655
- Adland E, Klenerman P, Goulder P, Matthews PC. Ongoing burden of disease and mortality from HIV/CMV coinfection in Africa in the antiretroviral therapy era. *Front Microbiol.* (2015) 6:1016. doi: 10.3389/fmicb.2015.01016
- Deayton JR, Sabin CA, Johnson MA, Emery VC, Wilson P, Griffiths PD. Importance of cytomegalovirus viraemia in risk of disease progression and death in HIV-infected patients receiving highly active antiretroviral therapy. *Lancet* (2004) 363:2116–21. doi: 10.1016/S0140-6736(04)16500-8
- Detels R, Leach CT, Hennessey K, Liu Z, Visscher BR, Cherry JD, et al. Persistent cytomegalovirus infection of semen increases risk of AIDS. *J Infect Dis.* (1994) 169:766–8. doi: 10.1093/infdis/169.4.766
- Griffiths PD. CMV as a cofactor enhancing progression of AIDS. *J Clin Virol.* (2006) 35:489–92. doi: 10.1016/j.jcv.2005.10.016
- McGoldrick SM, Bleakley ME, Guerrero A, Turtle CJ, Yamamoto TN, Pereira SE, et al. Cytomegalovirus-specific T cells are primed early after cord blood transplant but fail to control virus *in vivo*. *Blood* (2013) 121:2796–803. doi: 10.1182/blood-2012-09-453720
- Brown JA, Stevenson K, Kim HT, Cutler C, Ballen K, McDonough S, et al. Clearance of CMV viremia and survival after double umbilical cord blood transplantation in adults depends on reconstitution of thymopoiesis. *Blood* (2010) 115:4111–9. doi: 10.1182/blood-2009-09-244145
- Boppana SB, Britt WJ. Antiviral antibody responses and intrauterine transmission after primary maternal cytomegalovirus infection. *J Infect Dis.* (1995) 171:1115–21.
- Riddell SR, Watanabe KS, Goodrich JM, Li CR, Agha ME, Greenberg PD. Restoration of viral immunity in immunodeficient humans by the adoptive transfer of T cell clones. *Science* (1992) 257:238–41. doi: 10.1126/science.1352912
- Walter EA, Greenberg PD, Gilbert MJ, Finch RJ, Watanabe KS, Thomas ED, et al. Reconstitution of cellular immunity against cytomegalovirus in recipients of allogeneic bone marrow by transfer of T-cell clones from the donor. *N Engl J Med.* (1995) 333:1038–44.
- Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Ruchti F, et al. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *J Exp Med.* (2005) 202:673–85. doi: 10.1084/jem.20050882
- Karrer U, Sierro S, Wagner M, Oxenius A, Hengel H, Koszinowski UH, et al. Memory inflation: continuous accumulation of antiviral CD8+ T cells over time. *J Immunol.* (2003) 170:2022–9. doi: 10.4049/jimmunol.170.4.2022
- O'Hara GA, Welten SP, Klenerman P, Arens R. Memory T cell inflation: understanding cause and effect. *Trends Immunol.* (2012) 33:84–90. doi: 10.1016/j.it.2011.11.005
- Kovacs A, Schluchter M, Easley K, Demmler G, Shearer W, La Russa P, et al. Cytomegalovirus infection and HIV-1 disease progression in infants born to HIV-1-infected women. pediatric pulmonary and cardiovascular complications of vertically transmitted HIV infection study group. *N Engl J Med.* (1999) 341:77–84.
- Filteau S, Rowland-Jones S. Cytomegalovirus infection may contribute to the reduced immune function, growth, development, and health of HIV-exposed, uninfected african children. *Front Immunol.* (2016) 7:257. doi: 10.3389/fimmu.2016.00257
- Looney RJ, Falsey A, Campbell D, Torres A, Kolassa J, Brower C, et al. Role of cytomegalovirus in the T cell changes seen in elderly individuals. *Clin Immunol.* (1999) 90:213–9.
- Pawelec G. T-cell immunity in the aging human. *Haematologica* (2014) 99:795–7. doi: 10.3324/haematol.2013.094383
- Robins HS, Srivastava SK, Campregher PV, Turtle CJ, Andriesen J, Riddell SR, et al. Overlap and effective size of the human CD8+ T cell receptor repertoire. *Sci Transl Med.* (2010) 2:47ra64. doi: 10.1126/scitranslmed.3001442
- Chen G, Yang X, Ko A, Sun X, Gao M, Zhang Y, et al. Sequence and structural analyses reveal distinct and highly diverse human CD8(+) TCR repertoires to immunodominant viral antigens. *Cell Rep.* (2017) 19:569–83. doi: 10.1016/j.celrep.2017.03.072
- Yang X, Gao M, Chen G, Pierce BG, Lu J, Weng NP, et al. Structural basis for clonal diversity of the public T cell response to a dominant human cytomegalovirus epitope. *J Biol Chem.* (2015) 290:29106–19. doi: 10.1074/jbc.M115.691311
- Miles DJ, van der Sande M, Jeffries D, Kaye S, Ismaili J, Ojuola O, et al. Cytomegalovirus infection in Gambian infants leads to profound CD8 T-cell differentiation. *J Virol.* (2007) 81:5766–76. doi: 10.1128/JVI.00052-07
- Malik A, Adland E, Laker L, Kløverpris H, Fardos R, Roeder J, et al. Immunodominant cytomegalovirus-specific CD8+ T-cell responses in sub-Saharan African populations. *PLoS ONE* (2017) 12:e0189612. doi: 10.1371/journal.pone.0189612
- Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, et al. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HL. *Nature* (2004) 432:769–75. doi: 10.1038/nature03113
- Macdonald WA, Purcell AW, Mifsud NA, Ely LK, Williams DS, Chang L, et al. A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T cell recognition. *J Exp Med.* (2003) 198:679–91. doi: 10.1084/jem.20030066
- Payne R, Muenchhoff M, Mann J, Roberts HE, Matthews P, Adland E, et al. Impact of HLA-driven HIV adaptation on virulence in populations of high HIV seroprevalence. *Proc Natl Acad Sci USA.* (2014) 111:E5393–400. doi: 10.1073/pnas.1413339111
- Leisner C, Loeth N, Lamberth K, Justesen S, Sylvester-Hvid C, Schmidt EG, et al. One-pot, mix-and-read peptide-MHC tetramers. *PLoS ONE* (2008) 3:e1678. doi: 10.1371/journal.pone.0001678
- Rius C, Attaf M, Tungatt K, Bianchi V, Legut M, Bovay A, et al. Peptide-MHC class I tetramers can fail to detect relevant functional T cell clonotypes and underestimate antigen-reactive T cell populations. *J Immunol.* (2018) 200:2263–79. doi: 10.4049/jimmunol.1700242
- Emerson RO, DeWitt WS, Vignali M, Gravley J, Hu JK, Osborne EJ, et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat Genet.* (2017) 49:659–65. doi: 10.1038/ng.3822
- Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* (2015) 12:380–1. doi: 10.1038/nmeth.3364
- Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol.* (2015) 11:e1004503. doi: 10.1371/journal.pcbi.1004503
- Attaf M, Huseby E, Sewell AK. Alphabeta T cell receptors as predictors of health and disease. *Cell Mol Immunol.* (2015) 12:391–9. doi: 10.1038/cmi.2014.134
- Argaet VP, Schmidt CW, Burrows SR, Silins SL, Kurilla MG, Doolan DL, et al. Dominant selection of an invariant T cell antigen receptor in response to persistent infection by Epstein-Barr virus. *J Exp Med.* (1994) 180:2335–40. doi: 10.1084/jem.180.6.2335
- Lehner PJ, Wang EC, Moss PA, Williams S, Platt K, Friedman SM, et al. Human HLA-A0201-restricted cytotoxic T lymphocyte recognition of influenza A is dominated by T cells bearing the V beta 17 gene segment. *J Exp Med.* (1995) 181:79–91. doi: 10.1084/jem.181.1.79
- Moss PA, Moots RJ, Rosenberg WM, Rowland-Jones SJ, Bodmer HC, McMichael AJ, et al. Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. *Proc Natl Acad Sci USA.* (1991) 88:8987–90. doi: 10.1073/pnas.88.20.8987
- Kløverpris HN, McGregor R, McLaren JE, Ladell K, Harndahl M, Stryhn A, et al. CD8+ TCR bias and immunodominance in HIV-1 infection. *J Immunol.* (2015) 194:5329–45. doi: 10.4049/jimmunol.1400854
- Venturi V, Chin HY, Asher TE, Ladell K, Scheinberg P, Bornstein E, et al. TCR beta-chain sharing in human CD8+ T cell responses to cytomegalovirus and EBV. *Immunol V* (2008) 181:7853–62. doi: 10.4049/jimmunol.181.11.7853
- Venturi V, Chin HY, Price DA, Douek DC, Davenport MP. The role of production frequency in the sharing of simian immunodeficiency virus-specific CD8+ TCRs between macaques. *J Immunol.* (2008) 181:2597–609. doi: 10.4049/jimmunol.181.4.2597
- Quigley ME, Greenaway HY, Venturi V, Lindsay R, Quinn KM, Seder RA, et al. Convergent recombination shapes the clonotypic landscape of

- the naive T-cell repertoire. *Proc Natl Acad Sci USA*. (2010) 107:19414–9. doi: 10.1073/pnas.1010586107
39. Venturi V, Kedzierska K, Price DA, Doherty PC, Douek DC, Turner SJ, et al. Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination. *Proc Natl Acad Sci USA*. (2006) 103:18691–6. doi: 10.1073/pnas.0608907103
40. Valkenburg SA, Josephs TM, Clemens EB, Grant EJ, Nguyen TH, Wang GC, et al. Molecular basis for universal HLA-A\*0201-restricted CD8+ T-cell immunity against influenza viruses. *Proc Natl Acad Sci USA*. (2016) 113:4440–5. doi: 10.1073/pnas.1603106113
41. Miles JJ, Bulek AM, Cole DK, Gostick E, Schauenburg AJ, Dolton G, et al. Genetic and structural basis for selection of a ubiquitous T cell receptor deployed in Epstein-Barr virus infection. *PLoS Pathog*. (2010) 6:e1001198. doi: 10.1371/journal.ppat.1001198
42. Cole DK, Yuan F, Rizkallah PJ, Miles JJ, Gostick E, Price DA, et al. Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J Biol Chem*. (2009) 284:27281–9. doi: 10.1074/jbc.M109.022509
43. Bovay A, Zoete V, Dolton G, Bulek AM, Cole DK, Rizkallah PJ, et al. T cell receptor alpha variable 12-2 bias in the immunodominant response to Yellow fever virus. *Eur J Immunol*. (2018) 48:258–72. doi: 10.1002/eji.201747082
44. Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses? *Nat Rev Immunol*. (2008) 8:231–8. doi: 10.1038/nri2260
45. Venturi V, Quigley MF, Greenaway HY, Ng PC, Ende ZS, McIntosh T, et al. A mechanism for TCR sharing between T cell subsets and individuals revealed by pyrosequencing. *J Immunol*. (2011) 186:4285–94. doi: 10.4049/jimmunol.1003898
46. Attaf M, Holland SJ, Bartok I, Dyson J. Alphabeta T cell receptor germline CDR regions moderate contact with MHC ligands and regulate peptide cross-reactivity. *Sci Rep*. (2016) 6:35006. doi: 10.1038/srep35006
47. Furmanskı AL, Ferreira C, Bartok I, Dimakou S, Rice J, Stevenson FK, et al. Public T cell receptor beta-chains are not advantaged during positive selection. *J Immunol*. (2008) 180:1029–39. doi: 10.4049/jimmunol.180.2.1029
48. Naeger DM, Martin JN, Sinclair E, Hunt PW, Bangsberg DR, Hecht F, et al. Cytomegalovirus-specific T cells persist at very high levels during long-term antiretroviral treatment of HIV disease. *PLoS ONE* (2010) 5:e8886. doi: 10.1371/journal.pone.0008886
49. Sierro S, Rothkopf R, Klenerman P. Evolution of diverse antiviral CD8+ T cell populations after murine cytomegalovirus infection. *Eur J Immunol*. (2005) 35:1113–23. doi: 10.1002/eji.200425534
50. Trgovcich J, Kincaid M, Thomas A, Griessl M, Zimmerman P, Dwivedi V, et al. Cytomegalovirus reinfections stimulate CD8 T-memory inflation. *PLoS ONE* (2016) 11:e0167097. doi: 10.1371/journal.pone.0167097
51. Kim J, Kim AR, Shin EC. Cytomegalovirus infection and memory T cell inflation. *Immune Netw*. (2015) 15:186–90. doi: 10.4110/in.2015.15.4.186
52. Cicin-Sain L, Brien JD, Uhrlaub JL, Drabig A, Marandu TF, Nikolich-Zugich J. Cytomegalovirus infection impairs immune responses and accentuates T-cell pool changes observed in mice with aging. *PLoS Pathog*. (2012) 8:e1002849. doi: 10.1371/journal.ppat.1002849
53. Betts MR, Ambrozak DR, Douek DC, Bonhoeffer S, Brenchley JM, Casazza JP, et al. Analysis of total human immunodeficiency virus (HIV)-specific CD4(+) and CD8(+) T-cell responses: relationship to viral load in untreated HIV infection. *J Virol*. (2001) 75:11983–91. doi: 10.1128/JVI.75.24.11983-11991.2001
54. Welten SPM, Sandu I, Baumann NS, Oxenius A. Memory CD8 T cell inflation vs tissue-resident memory T cells: same patrollers, same controllers? *Immunol Rev*. (2018) 283:161–75. doi: 10.1111/imr.12649
55. Messaoudi I, Guevara Patiño JA, Dyall R, LeMaout J, Nikolich-Zugich J. Direct link between mhc polymorphism, T cell avidity, and diversity in immune defense. *Science* (2002) 298:1797–800. doi: 10.1126/science.1076064
56. Zinkernagel RM, Doherty PC. Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytic choriomeningitis. *Nature* (1974) 251:547–8. doi: 10.1038/251547a0
57. Carlson JM, Le AQ, Shahid A, Brumme ZL. HIV-1 adaptation to HLA: a window into virus-host immune interactions. *Trends Microbiol*. (2015) 23:212–24. doi: 10.1016/j.tim.2014.12.008

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer EH declared a past co-authorship with one of the authors AS to the handling editor.

Copyright © 2018 Attaf, Malik, Severinsen, Roider, Ogongo, Buus, Ndung'u, Leslie, Kløverpris, Matthews, Sewell and Goulder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.